Nonconvex Regularization for Image Segmentation

R. Chartrand
Theoretical Division
Los Alamos National Laboratory
Los Alamos, New Mexico, USA

V. Staneva
Theoretical Division
Los Alamos National Laboratory
Los Alamos, New Mexico, USA

Abstract - *We propose a new method for image segmentation based on a variational regularization algorithm for image denoising. We modify the Rudin-Osher-Fatemi (ROF) model in [1] by minimizing the L^p-norm of the gradient, where $p > 0$ is very small. The result is that we better preserve edges, while flattening regions away from the edges. This results in an automatic segmentation of the image into several regions, which does not require any prior knowledge about the number of those regions, or their intensity levels.*

Keywords: Image segmentation, image denoising, total variation regularization, nonconvex optimization.

1 Introduction

Two of the basic problems in image processing are denoising and segmentation. They are closely related, with similar objectives: given a noisy image, return a noise-free image while preserving important information from the original. Perhaps the greatest amount of information in an image is contained in the edges of objects. Consequently, preserving edges in an image is of paramount importance for both denoising and segmentation. One can regard the two problems as differing only in degree: in denoising, one seeks to remove noise and as little else as possible, while in segmentation, the goal is to remove all variation except for the edges of image regions.

Thus it is not surprising that there are models that are used for both problems. The best-known example of this is the Mumford-Shah functional [2]. A noisy image d defined on $\Omega \subset \mathbf{R}^2$ is denoised by computing the minimizer of the following functional:

$$F_{MS}(u,\Gamma) = \int_{\Omega\setminus\Gamma} |\nabla u|^2 + \nu|\Gamma| + \frac{\lambda}{2}\int_\Omega |u-d|^2. \qquad (1)$$

Here, Γ represents the union of all edges in the reconstruction u. The first term of (1) is a regularization term; it ensures that the denoised image will be smooth on each component of $\Omega \setminus \Gamma$. The second term penalizes the total length of the curves making up the edge set Γ, which keeps the noise of d from inducing spurious variation in the edges. The final term is a data fidelity term, which keeps the result close to the original image. The parameters ν and λ allow one to adjust the relative effect of the terms.

In [2], Mumford and Shah regarded the minimizer of F_{MS} as a segmentation of d, though it more closely resembles variational denoising methods that have been developed since. For segmentation, a more common approach is to assume that u in (1) is piecewise constant, an approach also considered in [2]. An interesting special case is where u takes on only two values, say 0 and 1. Then u is the characteristic function of a set E, and the perimeter of E is the total variation of u:

$$F_{ROF}(u) = \int_\Omega |\nabla u| + \frac{\lambda}{2}\int_\Omega |u-d|^2. \qquad (2)$$

Note that the regularization term of (1) is now zero, and we can subsume ν into λ. This functional is exactly that of the Rudin-Osher-Fatemi (ROF) model [1], a popular denoising method. We thus have that binary-value Mumford-Shah segmentation is the restriction to binary functions of Rudin-Osher-Fatemi denoising.

2 Total p-variation segmentation

In this work, we simultaneously modify the regularization of edges and the image away from the edges, by introducing a small exponent $p > 0$:

$$F_p(u) = \frac{1}{p}\int_\Omega |\nabla u|^p + \frac{\lambda}{2}\int_\Omega |u-d|^2. \qquad (3)$$

This has two effects. First, small but nonzero values of $|\nabla u|$ are penalized more, which makes the minimizer tend to be piecewise constant. We thus obtain segmentation without explicitly assuming a particular form for u. (This is the same phenomenon as the well known "staircasing" artifact of the ROF model, taken to a greater degree.) Second, the penalty on the length of edges is replaced by a penalty on a $(2-p)$-dimensional measure of the edges. To be specific, it is shown in [3] that if u is the characteristic function of E, then unless the boundary of E has upper box-counting dimension at least $2-p$, the discretization of $\int|\nabla u|^p$ tends to zero as the grid size tends to zero. In practice, given a positive grid size, the discretization of $\int|\nabla u|^p$ places some penalty on edges, but it is much weaker than that of total variation. This allows the segmentation to capture the boundaries of complicated regions more accurately.

3 Implementation

To compute a minimizer of (3), we use a straightforward generalization of the fixed-point method of Vogel and Oman [4]. Consider the Euler-Lagrange equation of $\min_u F_p(u)$:

$$0 = -\nabla \cdot \left(|\nabla u|^{p-2} \nabla u\right) + \lambda(u-d). \tag{4}$$

We solve (4) iteratively, by substituting the previous iterate u_{n-1} into $|\nabla u|$, then letting u_n be the solution of the resulting linear equation:

$$\left(-\nabla \cdot \left(|\nabla u_{n-1}|^{p-2} \nabla\right) + \lambda I\right) u_n = \lambda d. \tag{5}$$

Also, to avoid division by zero, we approximate $|\nabla u_{n-1}|$ by $\sqrt{|\nabla u_{n-1}|^2 + \varepsilon}$ for a small ε (namely 10^{-6} in the examples below). We typically begin the iteration with $u_0 = d$, the noisy image.

We discretize the problem with a uniform, rectangular grid with spacing Δx. We consider u and d to be in vectorized form: if the images are of size $m \times n$, then u and d are vectors of length $N = mn$. Let D_x, D_y be the matrices representing the finite-difference approximations of differentiation with respect to x and y. To be specific, we use forward differencing with Neumann boundary conditions. Thus, (5) takes the form

$$\left(D_x^T Q_{n-1} D_x + D_y^T Q_{n-1} D_y + \lambda I\right) u_n = \lambda d, \tag{6}$$

where Q_{n-1} is a diagonal matrix with entries

$$\left((D_x u_{n-1})_i^2 + (D_y u_{n-1})_i^2 + \varepsilon\right)^{(p-2)/2}. \tag{7}$$

We solve (6) with a preconditioned conjugate gradient method.

We do not prove convergence. Moreover, this procedure can only be expected to produce a local minimum, owing to the nonconvexity of (3). However, in hundreds of examples we have always observed convergence, unless ε is too small, in which case the system (6) can become numerically singular. We have no way of knowing whether the computed minimizer is local or global, but in our experience it is always a sensible reconstruction.

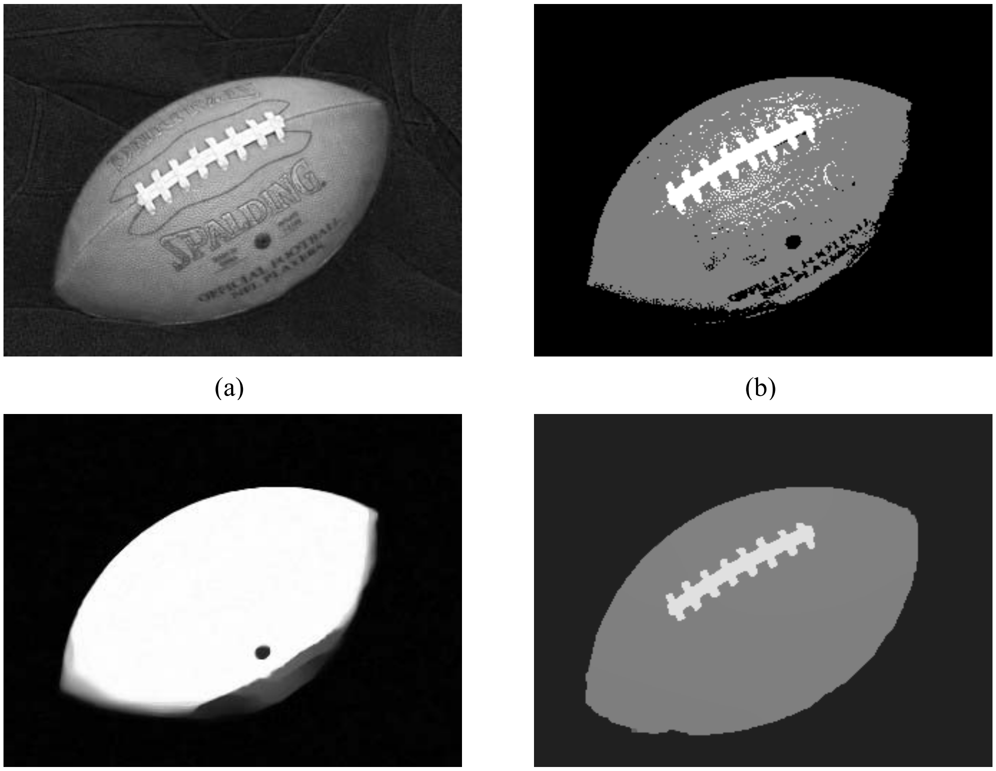
4 Results

First, we consider an image of a football (Figure 1(a)). It is not very difficult to separate the football from the background. However, it is a more challenging task to distinguish between the ball and the stitching, and reconstruct its edges correctly. First we apply popular algorithms for image segmentation. *k*-means is an algorithm which partitions the intensity values of the image into *k* clusters. We choose $k = 3$, and achieve the segmentation shown in Figure 1(b). The noise in the image results in rough boundaries. Another algorithm, which uses total variation regularization to eliminate noise, is the convexified version of the Chan-Vese level set method [5] proposed in [6]. Its disadvantage, though, is that it is designed to obtain binary segmentation; we see in Figure 1(c) that the third region, the stitching, has been lost. Then we apply our *p*-TV regularization with $p = 0.01$. The result achieved after 20 iterations is displayed in Figure 1(d). The noise does not affect the segmentation because the regularization term in the functional accounts for it, while at the same time, preserves the edges of the different regions.

For our second example, we use the well-known cameraman image. The intensity values range from 0 to 1, and we add Gaussian noise of standard deviation 0.1 (Figure 2(a)). We again have $p = 0.01$. For an initial guess we first use the noisy image. The result we obtain after 20 iterations is displayed in Figure 2(c). Although the initial image is a complex grayscale image, which even without noise is a difficult segmentation problem, the final image consists of only a few grayscale levels. The number of the final regions and the crudeness of their boundaries can be regulated through the regularization parameter λ. If we want the final image to be close to a binary image, we can implement the algorithm using as a starting point an image obtained by thresholding the noisy image. The result after 10 iterations has even fewer grayscale levels (Figure 2(e)). The observation that different starting points lead to different final results shows the existence of local minima. This allows us to vary the segmentation of the image, which will reflect properties of the initial guess.

We remark that the algorithm performs essentially the same with $p = 0$. On the one hand, the iteration (5) makes sense in this case; on the other hand, it is not solving an Euler-Lagrange equation, as the functional F_p we are trying to minimize is undefined when $p = 0$.

The performance of the method does not depend exclusively on having an L^2 norm as a data fidelity term. For



(a) (b) (c) (d)

Figure 1: (a) An image of a football: the task is to find three separate regions in it; (b) segmentation obtained by the *k*-means algorithm: the regions are corrupted by noise; (c) result obtained by a total variation regularized segmentation: only two regions are defined; (d) segmentation achieved by total *p*-variation regularization with a small *p*: the image is segmented in three regions with clear, accurate edges.

example, we can implement it with an L^1 norm, known for its ability to remove salt-and-pepper noise. On the cameraman image in Figure 2(b), in which 10% of the pixels have been corrupted by salt and pepper noise, the algorithm exhibits the same behavior. Results with two different starting points are shown in Figures 2(d) and 2(f). A Poisson noise data fidelity term [7] can also be fit into the model. This flexibility of the algorithm makes it applicable to a wide variety of segmentation problems.

5 References

[1] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, pp. 259–268, 1992.

[2] D. Mumford and J. Shah, “Optimal approximations by piecewise smooth functions and associated variational problems,” *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.

[3] R. Chartrand, “Nonconvex regularization for shape preservation,” in *2007 IEEE International Conference on Image Processing*, to appear.

[4] C. R. Vogel and M. E. Oman, “Iterative methods for total variation denoising,” *SIAM J. Sci. Comput.*, vol. 17, pp. 227–238, 1996.

[5] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Trans. Image Process.*, vol. 10, pp. 266–277, 2001.

[6] T. F. Chan, S. Esedoḡlu, and M. Nikolova, “Algorithms for finding global minimizers of image segmentation and denoising models,” *SIAM J. Appl. Math.*, vol. 66, pp. 1632–1648, 2006.

[7] T. Le, R. Chartrand, and T. J. Asaki, “A variational approach to reconstructing images corrupted by Poisson noise,” *J. Math. Imaging Vision*, 2007. To appear.

Figure 2: (a) The cameraman image corrupted by Gaussian noise; (b) the cameraman image corrupted by salt and pepper noise; (c) segmentation of (a) obtained with the noisy image as an initial guess; (d) segmentation of (b) obtained with the noisy image as an initial guess; (e) segmentation of (a) obtained with the thresholded image as an initial guess; (f) segmentation of (b) obtained with the thresholded image as an initial guess.